EPA
United States
Environmental Protection
Agency

# Best Practices for Identifying Reference Condition in Mid-Atlantic Streams

## Reference Condition Concept

The Clean Water Act (CWA) poses significant challenges to states and tribes charged with evaluating whether aquatic resources under their management achieve the biological integrity objective and the "protection and propagation" goals. One of the critical challenges is the development of a standard or benchmark by which to judge whether particular water bodies are in accordance with the CWA objective and goals. The concept of a reference condition and its implementation form the foundation on which to make such judgments (Stoddard et al. 2006a)

Reference conditions have been applied at site-specific and regional scales. Regional reference condition, described here, is recommended to support biological criteria. Biological criteria are used to detect deviation from reference condition to determine whether water bodies meet their water quality standards. The Biological Condition Gradient (BCG) is a scientific narrative model for interpreting biological response to increasing effects of disturbance on aquatic ecosystems. The BCG describes how attributes of aquatic ecosystems change in response to increasing levels of human disturbance (Fig. 1, Davies and Jackson 2006).

States in the Mid-Atlantic region have developed and implemented the concept of reference condition in a variety of ways to meet their individual needs, without comprehensive guidance from EPA. This brochure offers examples from these states as case studies in the application of the reference condition concept in water resource management.

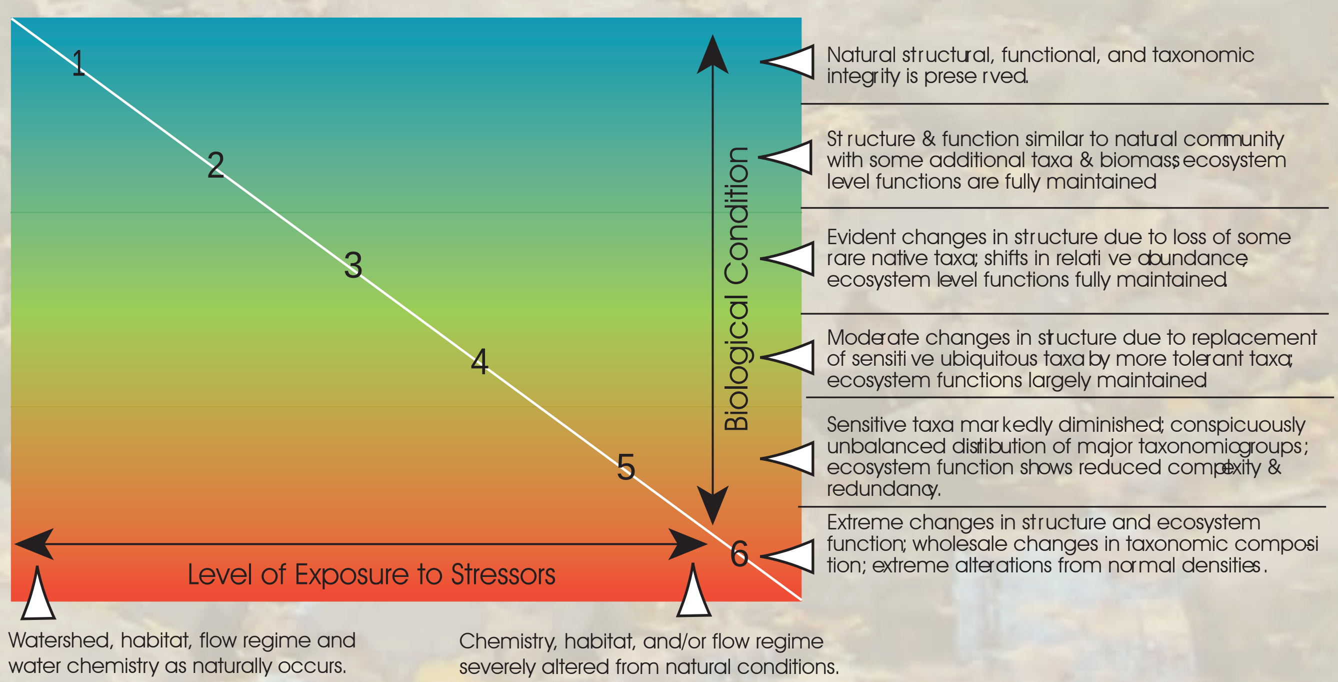


**Figure 1.** The Biological Condition Gradient (BCG) is a tool for developing more precise aquatic life uses. The BCG displays six positions of biological condition along a stressor-response curve, with Position 1 exhibiting the least stress and highest quality condition, and Position 6 representing the greatest stress and lowest quality. (Modified from EPA 2005)

## Types of Reference Condition

The term reference condition can have multiple meanings. Therefore, consistent and specified definitions of reference condition can greatly enhance collaboration among states and the transfer of best practice technology and expertise. In all cases, reference condition is the benchmark against which changes in current biological conditions are evaluated. The following definitions distinguish among four specified types of reference condition:

Reference Condition for Biological Integrity, RC(BI): the natural biological condition of a water body, undisturbed by human activity. As a conceptual aid, it is useful to think of an absolute "natural" or pristine condition that could exist in the absence of all historical and current human disturbances. This definition recognizes the need for a reference condition term reserved for "naturalness" or "biological integrity" even though we might only approximate it in most parts of the world.

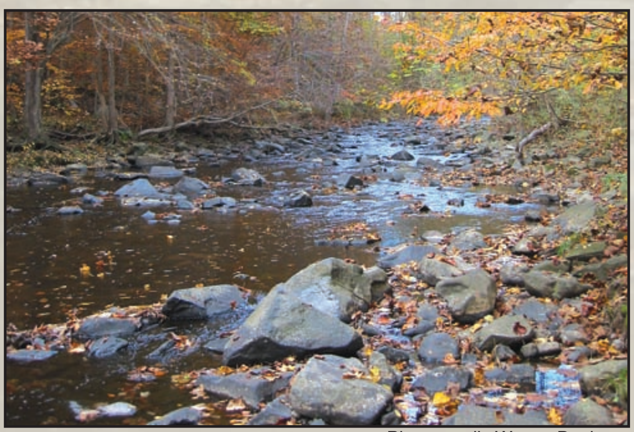

Photo credit: Wayne Davis

**Figure 2.** A Mid-Atlantic stream displaying physical attributes suggestive of minimally disturbed condition.

Minimally Disturbed Condition, MDC: the biological condition found in water bodies in landscapes with minimum human disturbance. Places that meet the criteria for RC(BI) are rare or impossible to find. Human activity is present throughout the global ecosystem, affecting remote systems through long-range atmospheric transport and deposition of pollutants onto pristine landscapes. Biological integrity in places with a low amount of human activity might not be significantly affected. The phrase "minimally disturbed condition" describes the biological condition in places with a minimal amount of human disturbance.

Least Disturbed Condition, LDC: in altered landscapes, the biological condition found in water bodies with the least amount of human disturbance compared to similar water bodies in the region of interest. There is a clear need to describe the best condition of water bodies in landscapes that have been moderately to heavily disturbed by human activities. Definitions like "minimally disturbed condition" are of little practical use in these situations. Therefore, the phrase "least disturbed condition" has been applied to describe the condition in water bodies that are the least disturbed in a landscape altered by significant human activity. LDC should not be used as a benchmark for biological integrity. Further, in certain severely altered landscapes, LDC may not even be useful as a benchmark for meeting CWA aquatic life use protection and propagation goals.

Best Potential Condition, BPC: the highest possible biological condition deemed achievable through the implementation of best management practices and other rehabilitation activities that can be undertaken in a given landscape given social and economic considerations. In some circumstances, a condition could be achieved that is better than the least disturbed condition (i.e., better than the condition at the best sites) with implementation of the best available practices to remove or minimize stressors. Even though the biological potential might approach biological integrity if the stressors are removed, societal/economic constraints typically mean that a condition is achieved that differs from biological integrity. The term "best potential condition" describes this condition, where the biological expectations are set somewhere between the least disturbed condition and biological integrity.

## Approaches to Deriving Reference Condition

Ideally, every water body segment will have its own reference condition. In order to describe reference condition in a way that will not change over time due to further human activity, several general methods have been developed through practical experience. These methods include sampling the biota at sites with little or no indication of stressors associated with human disturbance (i.e., minimally disturbed reference sites). In altered landscapes, where such sites are few or absent, reference conditions are determined through a combination of methods: (1) sampling biota from least disturbed sites (reference sites), (2) interpreting historical records to deduce which biological characteristics occurred at times with substantially less human disturbance, (3) developing models that incorporate the best ecological knowledge, and (4) using best professional judgment.

1. Reference Sites
The selection and characterization of reference sites that are minimally or not disturbed by human activities have been the basis for defining reference conditions that approximate biological integrity. However, reference sites have also been identified as "the best of what's left" and as such are used to estimate a least disturbed condition. The approach used to select reference sites may be similar, regardless of whether the sites are classified as minimally disturbed or least disturbed. The process and considerations involved in selecting reference sites is discussed in further detail in the following pages.

2. Historical Reconstruction
The role of historical reconstruction is to use available data to describe a range of water body or riparian conditions that existed at an earlier time. Historical reconstruction estimates a minimally disturbed condition rather than a least disturbed condition. Benefits of this approach include the following:

- Improving the characterization obtained from reference sites;
- Needing to be generated only once;
- Providing a permanent benchmark;
- Allowing for a more cost-effective approach than extensive sampling; and,
- Providing motivation to stakeholders as a vision of desirable conditions.

3. Empirical modeling
When the number of representative reference sites is low and historical information is not sufficient to reconstruct reference condition, predictive modeling can be used to construct and calibrate a model reference condition. This approach effectively leverages a smaller number of sites from the region or water body type than is needed for the typical spatially intensive reference site approach. However, it does require reliable data from representative sites. Absent such data, this approach reverts to a best professional judgment approach with its inherent shortcomings of subjectivity. This approach is limited also by the data used in the model, and therefore, inferences beyond those data must be undertaken with great caution.

Predictive modeling approaches that show promise include the following:

- Extension of reference site results from adjacent regions or similar water body types;
- Application of stressor-biotic assemblage interactions identified in restoration experiments; and,
- Inclusion of extirpated species or exclusion of nonnative species.

Photo credit: Morris Perot

**Figure 3.** Biologists conducting a physical habitat assessment at an unnamed tributary in Frederick County, MD.

4. Best Professional Judgment
Best professional judgment (BPJ) should be incorporated into all decisions; however, comparisons with a stressor gradient are also needed to accurately identify the position of the reference sites. The natural variability in biological attributes for a certain biological condition (e.g., biological integrity) dictates that biological data alone should not be used to develop reference conditions. Although BPJ is a critical part of biological assessments, great care must be taken to ensure that the development of reference conditions and selection of reference sites are well documented and include objective procedures that can be reproduced easily by others.

## Selecting Reference Sites

Selecting reference sites involves applying screening criteria to a set of sites from the region or water body type of interest. These criteria will identify sites that are most likely to be minimally or least disturbed by human activities. The result is a set of *candidate* reference sites, qualified as such because depending upon the protectiveness that the screening criteria represents, some of the sites selected may not represent the desired reference condition (Fig. 4).

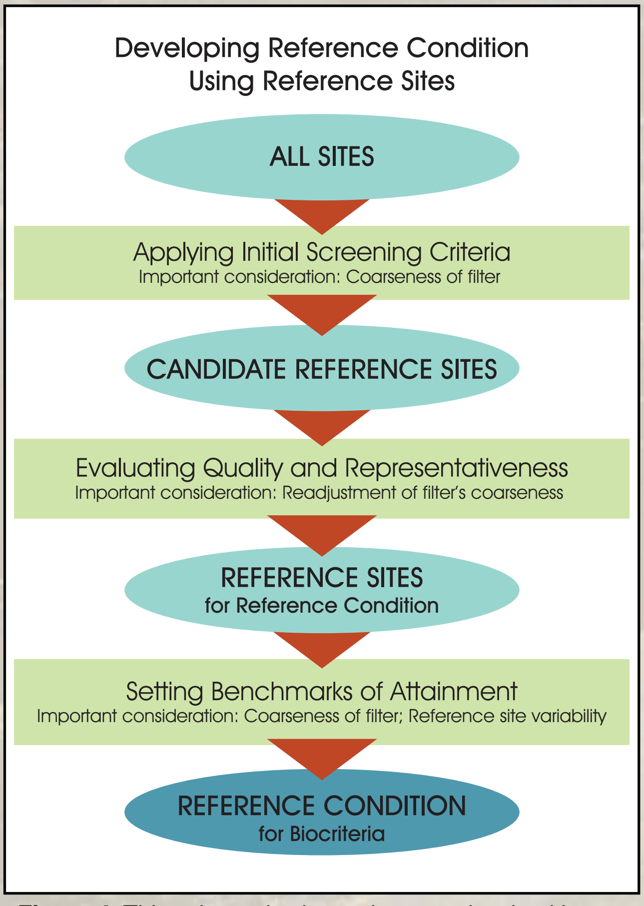


**Figure 4.** This schematic shows the steps involved in developing reference condition using reference sites. At each step, it is important to consider the coarseness of the filter, which influences the confidence that the quality of the reference sites accurately represents the intended quality of the reference condition.

Criteria useful for screening sites include stressors, indicators of stressor sources, and indicators along the pathway from source to exposure. The goal is to evaluate as many of these indicators as is practical, efficient, and relevant to the water body type and region of interest. Only those sites that meet the criteria of minimal disturbance or those that are the least disturbed among the set of sites are considered candidate reference sites.

Applying Reference Criteria
Screening criteria should apply to potential reference sites at three levels: landscape or watershed, reach or riparian corridor, and site. While watershed and reach scale analysis using maps and/or GIS technology can identify candidate reference areas, there are many human activities that can only be revealed by collecting data at the site level. However, it is essential that the final list of candidate sites identified through this screening be evaluated against an objective set of stressor thresholds obtained from sampling data. Brief descriptions of the Mid-Atlantic Highlands Streams Assessment and the Maryland Biological Stream Survey illustrate approaches to reference site selection using objective criteria.

Having too few reference sites to analyze each region or water body type is a common constraint in indicator development and may lead to relaxation of reference site criteria to obtain more sites.

Prior to the ***Mid-Atlantic Highlands Streams Assessment*** (***MAHSA***; Davis and Scott 2000), a precedent for using objective, abiotic criteria for identifying reference sites in the region had not yet been set. MAHSA biologists anticipated that the probabililty-based design would generate too few candidate reference sites, and therefore asked local resource managers to "hand-pick" sites in order to augment the spectrum of conditions. All sites sampled--including those selected via BPJ--were subjected to objective, abiotic criteria (below) representing potential disturbance activities including acid rain, acid mine drainage, agriculture, and general development. Sites were considered candidate reference sites if they met these criteria:

- acid neutralizing capacity (ANC) > 50 µeq/L (pH of about 6)
- sulfate ($SO_4$) < 400 µeq/L
- total phosphorus (P) < 20 µg/L
- total nitrogen (N) < 750 µg/L
- chloride ($Cl^-$) <100 µeq/L
- mean Rapid Bioassessment Protocol (RBP) habitat score > 15

Application of the objective screening filter to the 58 BPJ sites resulted in only 15 sites (26%) meeting the reference site criteria. This study established the need to support the application of objective, abiotic screening criteria for candidate reference sites regardless of the method used to select sites.

The ***Maryland Biological Stream Survey (MBSS)*** is an example of a probability-based stream monitoring program that successfully uses objective criteria for sampled sites to identify reference conditions (Klauda et al. 1998). Because the program samples approximately 300 sites per year, data are available from all regions and stream types. No screening was done to narrow the range of candidate reference sites, so all sampled sites were evaluated using the following criteria of water chemistry, physical habitat, and land use stress (Roth et al. 1998). To develop the Maryland fish index of biological integrity (IBI), the MBSS used these criteria to identify 152 of the total 1098 sites (13.8%) as reference sites:

- pH ≥6 or blackwater stream (pH < 6 and DOC ≥8 mg/L)
- ANC ≥50 µeq/L
- dissolved oxygen ≥4 ppm
- nitrate ≤300 µeq/L (4.2 mg/L)
- urban land use ≤20% of the catchment area (draining to the site)
- forest land use ≥25% of the catchment area
- remoteness rating = optimal or suboptimal (>10 on 0-20 scale)
- aesthetics rating = optimal or suboptimal
- instream habitat rating = optimal or suboptimal
- riparian buffer width ≥15 m
- no channelization
- no point source discharges

The development of a fish IBI for MAHSA (Davis and Scott 2000) addressed this situation by evaluating metric performance against three different reference definitions: (1) least restrictive criteria based on chemical thresholds and the mean RBP habitat score (producing 46 reference sites with good geographic coverage); (2) moderately restrictive criteria based on chemical criteria, watershed land use, road density, and quantitative habitat filters (producing 23 reference sites with good geographic coverage); and (3) most restrictive criteria based on the moderately restrictive criteria plus the watershed condition class (Bryce et al. 1999; producing 12 reference sites with limited geographic coverage). If less restrictive criteria are used to define reference sites, then the lower quality may result in conditions that correspond to lower positions along the Biological Condition Gradient (e.g., Position 3 rather than 2, in Fig. 1).

Evaluating Quality and Representativeness

Because reference criteria vary among water quality monitoring programs, it is essential to evaluate the selected reference sites for whether they are truly minimally disturbed and whether they are representative of the water bodies of interest.

The Maryland Biological Stream Survey used a set of water chemistry, physical habitat, and land use reference criteria that produced reference sites with considerably higher benthic macroinvertebrate IBIs than all sites sampled and were also higher than the original MBSS set of reference sites (see dashed line in Fig. 5). Subsequently, the reference criteria used in the original MBSS IBIs were reviewed to identify changes that would result in greater confidence that the new reference sites could be defined as "minimally disturbed." Based on analysis of urban effects on stream condition (Vølstad et al. 2003), the presence of original reference sites with relatively high levels of urban land (i.e., 5% to 20%) indicated that not all reference sites were minimally disturbed; instead, many were impaired. Therefore, the MBSS changed the minimum allowable forested land use from ≥25% to ≥35% of the catchment area, maximum allowable urban land use from ≤20% to ≤5%, and minimum allowable riparian buffer from 15m to 30m. These changes in land use and riparian width thresholds resulted in a smaller proportion of stream sites meeting the reference site criteria. Using the original reference site criteria, 152 of the 1098 Round One sites (14%) were designated as reference sites. Using the new criteria, 196 of the total 2508 sites (8%) were designated as reference. These new reference sites were of higher quality (dotted line in Fig. 5) than sites meeting the original reference criteria. This result is consistent with greater confidence that the sites are minimally disturbed.

As the criteria for selecting reference sites are tightened to ensure minimally disturbed sites are chosen, it is important that these sites still represent the range of natural factors likely to control the biotic composition in the region of interest. For example, stream size, gradient, and elevation can be

## Setting Benchmarks for Reference Condition

The most common approach for using reference sites to set attainment benchmarks is to select a percentile of the reference site index scores below which is considered degraded. As previously discussed, the selection of the percentile should consider the coarseness, or restrictiveness, of the screening criteria. In the Mid-Atlantic Region, West Virginia sets a benchmark at the 5th percentile of the reference sites; Maryland uses the 9th percentile; Virginia uses the 10th percentile; and EPA's Mid-Atlantic Integrated Assessment (MAIA) uses the 25th percentile. In addition to percentile selection, the West Virginia Department of Environmental Protection (WVDEP), the Maryland Department of the Environment (MDE), and the Virginia Department of Environmental Quality (VDEQ) also include separate measures of index precision surrounding the benchmark in order to capture uncertainty for regulatory program use. These programs demonstrate innovative approaches for setting attainment benchmarks, as well as recommended practices for defining and implementing reference condition.

**WVSCI Scoring Criteria**

| |
|---|
| > 68.0<br>Unimpaired |
| ≥ 60.6 to 68<br>"Gray Zone" |
| < 60.6<br>Impaired |

**Figure 6.** The West Virginia Stream Condition Index (WVSCI) has a precision estimate of 7.4 units. Therefore, setting the benchmark at a score of 68 (the 5th percentile) effectively creates a "gray zone" between 68 and 60.6. (WVDEP 2006)

West Virginia Benchmark

West Virginia DEP set a low benchmark (at the 5th percentile) because their screening criteria included a secondary review of the candidate reference sites that passed the initial objective screening criteria (akin to the evaluation step in Fig. 4). Using BPJ in the secondary review, the number of sites dropped from 349 candidate reference sites to 216 reference sites (Southerland 2006). The BPJ included a review of each candidate site for its proximity to upstream point source discharges and an evaluation of anthropogenic activities and disturbances near the candidate sites. For their stream condition index (SCI), WVDEP (2006) determined a precision estimate of 7.4 out of 100 units. This uncertainty in the index itself sets the 5th percentile benchmark of 68 effectively to 60.6, the range between those values being in a "gray zone" (Fig. 6).

major determinants of potential biotic composition. Therefore the set of reference sites should cover and be limited to the range of natural conditions important in the region of interest. If the streams represent an elevation gradient, reference sites should also represent that elevation gradient; if streams considered represent a range of sizes, the set of reference sites should also represent these sizes.

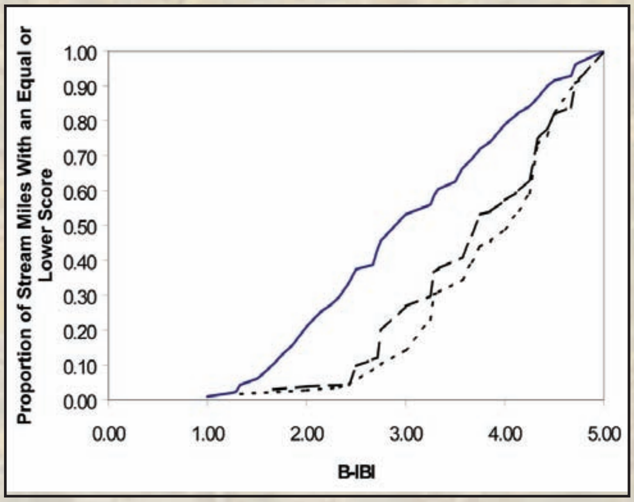


**Figure 5.** Cumulative distribution of stream miles with benthic macroinvertebrate IBI scores for (1) MBSS sites sampled in 2000-2004 (solid line), (2) subset of 2000-2004 sites meeting original reference criteria (dashed line), and (3) subset of 2000-2004 sites meeting new reference criteria (dotted line). (Southerland et al. 2005)

## Virginia Benchmark

Virginia DEQ (2006) sets the attainment benchmark at the 10th percentile of their reference sites. Much like West Virginia, the precision estimate for Virginia's stream condition index (VSCI) is ±7.9, which generates a "gray zone" around the benchmark. As a result, the 10th percentile benchmark of 60 effectively becomes a range from 55 to 63, where sites scoring above 63 are viewed as healthy streams and those below 55 are deemed moderately to severely stressed.

## Maryland Benchmark

The Maryland Department of Natural Resources did not establish a specific benchmark based on a distribution of their reference sites. They maintain a categorical index score rating similar to the metric scoring procedure of 1, 3 and 5, where any score above 3 is considered acceptable. To facilitate the use of the Maryland IBIs for the regulatory agency, a statistical measure of uncertainty (confidence interval) is used to determine whether the mean of the results from the sites sampled in a watershed is above or below the Index of Biotic Integrity (IBI) value considered indicative of satisfactory water quality (i.e., 3). Where at least ten sites have been sampled in a watershed, watershed-specific confidence intervals are calculated. If the upper bound of the confidence interval is less than 3, that watershed is designated as not meeting water quality criteria (MDE 2004). For comparison with other states' methods, it was determined that the MDE index benchmark of 3 was equivalent to the 9th percentile of the reference sites (Southerland 2006).

## Mid-Atlantic Integrated Assessment Benchmark

For a report card on the state of streams and rivers in the Mid-Atlantic region, data was combined from two sample surveys conducted in the region by MAIA from 1993 to 1998 (Stoddard et al. 2006b). MAIA established two benchmarks based on the distribution of reference sites. The 25th percentile value set the lower limit on "good" condition. The 1st percentile was used as the threshold below which values were deemed "poor." Values between the 1st and 25th percentiles were designated as "marginal." These classifications were deliberately used so as not to conflict with regulatory terms used by the States (Fig. 7).

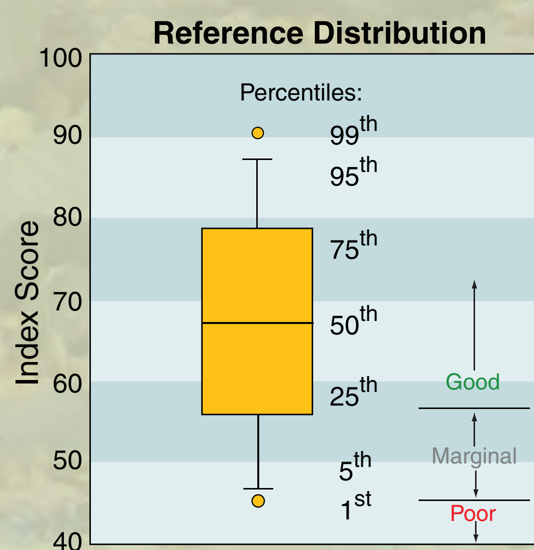


**Figure 7.** The two benchmarks set by MAIA established three categories of condition: poor, marginal and good. (Stoddard et al. 2006b)

Office of Environmental Information
Washington DC 20460
EPA-260-F-06-002
August 2006

## Contacts

Office of Environmental Information
Wayne Davis
davis.wayne@epa.gov

Office of Water
Evan Hornig
hornig.evan@epa.gov

Office of Research and Development
John Stoddard
stoddard.john@epa.gov

## References

Bryce, S.A., D.P. Larsen, R.M. Hughes, and P.R. Kaufmann. 1999. Assessing relative risks to aquatic ecosystems: A Mid-Appalachian case study. Journal of American Water Resources Association. 35:23-36.

Davies, S.P. and S.K. Jackson. 2006. The biological condition gradient: a descriptive model for interpreting change in aquatic ecosystems. Ecological Applications 16(4):1251-1266

Davis, W.S., and J. Scott. 2000. Mid-Atlantic Highlands Streams Assessment: Technical support document. EPA-903-B-00-004. U.S. Environmental Protection Agency, Region 3, Office of Research and Development, Ft. Meade, MD.

Klauda, R., P. Kazyak, S. Stranko, M. Southerland, N. Roth, and J. Chaillou. 1998. The Maryland Biological Stream Survey: A state agency program to assess the impact of anthropogenic stresses on stream habitat quality and biota. Third EMAP Symposium, Albany, NY. Environmental Monitoring and Assessment 51:299-316.

Maryland Department of the Environment (MDE). 2004. 2004 List of impaired surface waters [303(d) List] and integrated assessment of water quality in Maryland. Appendix C Listing Methodologies. Maryland Department of the Environment, Baltimore, MD. Available at http://www.mde.state.md.us/assets/document/AppndxC2004-303d_Final.pdf

Roth, N.E., M.T. Southerland, J.C. Chaillou, R.J. Klauda, P.F. Kazyak, S.A. Stranko, S.B. Weisberg, L.W. Hall, Jr., and R.P. Morgan II. 1998. Maryland Biological Stream Survey: Development of a fish index of biotic integrity. Environmental Management and Assessment 51:89-106.

Stoddard, J.L., D.P. Larsen, C.P. Hawkins, R.K. Johnson and R.H. Norris. 2006a. Setting expectations for the ecological condition of streams: the concept of reference condition. Ecological Applications 16(4):1267-1276.

Stoddard, J.L., A.T. Herlihy, B.H. Hill, R.M. Hughes, P.R. Kaufmann, D.J. Klemm, J.M. Lazorchak, F.H. McCormick, D.V. Peck, S.G. Paulsen, A.R. Olsen, D.P. Larsen, J. Van Sickle, and T.R. Whittier. 2006b. Mid-Atlantic Integrated Assessment (MAIA): State of the flowing waters report. EPA-620-12-06-001. U.S. Environmental Protection Agency, Office of Research and Development, Washington, DC. February 2006.

Southerland, M.T., G.M. Rogers, M.J. Kline, R.P. Morgan, D.M. Boward, P.F. Kazyak, R.J. Klauda, and S.A. Stranko. 2005. New biological indicators to better assess the condition of Maryland streams. CBWP-MANTA-EA-05-13. Report for the Maryland Department of Natural Resources, Annapolis, MD.

Southerland, M., J.H. Vølstad, L. Erb, E. Weber, and G. Rogers. 2006. Proof of concept for integrating bioassessment results from three state probabilistic monitoring programs. EPA-903-R-05-003. U.S. Environmental Protection Agency, Office of Environmental Information and Mid-Atlantic Integrated Assessment Program, Ft. Meade, MD.

U.S. Environmental Protection Agency (EPA). 2005. Use of biological information to better define designated aquatic life uses in state and tribal WQS: Tiered aquatic life uses. EPA-822-R-05-001. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

Virginia Department of Environmental Quality (VDEQ). 2006. Using probabilistic monitoring data to validate the Non-Coastal Virginia stream condition index. VDEQ Technical Bulletin WQA/2006-001. Richmond, VA. Available at: http://www.deq.virginia.gov/probmon/pdf/scival.pdf

Vølstad, J.H., N.E. Roth, M.T. Southerland, and G. Mercurio. 2003. Pilot study for Montgomery County and Maryland DNR data integration: Comparison of benthic macroinvertebrate sampling protocols for freshwater streams. EPA-903-R-03-005. U.S. Environmental Protection Agency, Office of Environmental Information and Mid-Atlantic Integrated Assessment, Fort Meade, MD.

West Virginia Department of Environmental Protection (WVDEP). 2006. West Virginia 2006 Section 303(d) List, updated 3/22/2006. Available at http://www.dep.state.wv.us/show_blob.cfm?ID=10203&Name=303(d)_Listing_Rationale_Only.pdf

Recycled /Recyclable
Printed with Vegetable Oil Based Inks on 100%
Post-consumer Process Chlorine Free Recycled Paper